2012•2013

FACULTY OF SCIENCES

Master of Statistics: Biostatistics

Masterproef

Validation of PFS as surrogate for overall survival for GBM

Promotor :
Prof. dr. Tomasz BURZYKOWSKI

Promotor :
Prof. THIERRY GORLIA

Julius Nangosyah

Master Thesis nominated to obtain the degree of Master of Statistics , specialization Biostatistics

Transnational University Limburg is a unique collaboration of two universities in two countries: the University of Hasselt and Maastricht University.

universiteit hasselt
KNOWLEDGE IN ACTION

Universiteit Hasselt | Campus Hasselt | Martelarenlaan 42 | BE-3500 Hasselt
Universiteit Hasselt | Campus Diepenbeek | Agoralaan Gebouw D | BE-3590 Diepenbeek

universiteit hasselt | Maastricht University

2012•2013
FACULTY OF SCIENCES
Master of Statistics: Biostatistics

Masterproef

Validation of PFS as surrogate for overall survival for GBM

Promotor :
Prof. dr. Tomasz BURZYKOWSKI

Promotor :
Prof. THIERRY GORLIA

Julius Nangosyah

Master Thesis nominated to obtain the degree of Master of Statistics , specialization Biostatistics

universiteit hasselt | Maastricht University

Certification

This is to certify that this report was written by NANGOSYAH JULIUS under our supervision.

Signature: _____ Date: _____

Prof. dr. Tomasz Burzykowski **Internal Supervisor**

Signature: _____ Date: _____

Prof. Thierry Gorlia **External Supervisor**

Thesis submitted in partial fulfilment of the requirements for the degree of Master of Statistics: Biostatistics.

Signature: _____ Date: _____

NANGOSYAH JULIUS **Student**

Acknowledgements

I wish to thank several people who have contributed to the successful realization of this project. First and foremost, I would like to thank European Organization for Research and Treatment of Cancer (EORTC) for giving me the opportunity to carry out this research. My sincere appreciation goes to my supervisors, Prof. dr. Tomasz Burzykowski and Thierry Gorlia for taking their time off to assist and guide me in doing this project. I also wish to extend my deepest appreciation to all my lecturers at Hasselt University under the program Master of Statistics with specialization in Biostatistics. It was a great honor for me to have attended this program. All the courses and advices have provided me a deeper understanding of the project. To all the University Staff, especially Mrs. Martine Machiels who helped us to settle, my heartfelt thanks.

Special thanks to all my classmates for all the intellectual enrichment, interesting, and fruitful discussions that I had with them. I am indebted to them for making my school experience a very memorable one! To all my group members; Geraldine, Josephine, and Susan from I have learnt a lot from you.

Finally, I wish to thank my family whose support has helped me to get this far. I am deeply thankful to my father, Mr. Nangosyah William for the guidance and support and to all siblings for all the love and trust. I am grateful to Merel Vergauwen and her family for their daily support and encouragement, particularly during the final days of writing my thesis. Above all, I am grateful to Almighty God for the good health and strength bestowed upon me for making all things possible. Indeed, with God nothing is impossible!

Julius Nangosyah
University of Hasselt
Belgium, September, 2013

Abstract

Background: Surrogate end points have great potential for use in clinical oncology, but their statistical validation presents major challenges[12]. A putative surrogate endpoint must be validated at both individual-level and trial-level before it can be used to replace the clinical endpoint in a future clinical trial[34]. Validated and approved surrogate endpoints are useful in development of new drugs. They save time and money since it takes a shorter time to give results. They also require fewer patients in the clinical trial.

Objective: To evaluate progression free survival as a surrogate for overall survival for the patients who have been newly diagnosed with Glioblastoma.

Methodology: Individual patients were available from two trials comparing radiotherapy alone (379 patients) with radiotherapy plus chemotherapy (389 patients).Correlation coefficients were estimated between the endpoints of PFS and OS through Cox proportional model and copula models. These methods were appropriate since both endpoints were failure time endpoints. Furthermore surrogate threshold effect approach was also applied.

Results: In the two trials, 778 patients (723)events were observed on OS while (746) events were observed on PFS. The rank correlation between PFS and OS was equal 0.845.The trial level surrogacy estimates from Cox proportional hazard ranged from 0.453 (0.36, 0.86) for unweighted regression to 0.429 (0.062, 0.852) for the weighted regression. Plackett copula estimates ranged from 0.48 (0.0746, 0.888) for unadjusted model to 0.896 (0.559, 1.233) for adjusted R-squared. The individual level surrogacy estimate from Plackett copula was 0.642 (0.635, 0.64). The hazard ratio of 0.135 or lower in terms of PFS would predict a non-zero treatment effect in terms of OS.

Conclusion: In this project PFS does not conveys clinically useful surrogacy properties for OS for patients with newly diagnosed Glioblastoma.PFS is not a valid surrogate for in Glioblastoma.

Contents

Acknowledgements . . . ii
Abstract . . . iii

1 Introduction **1**
1.1 Objectives . . . 2

2 Data Description **3**
2.1 Trial data description . . . 3
2.2 Combined data description . . . 3

3 Methodology **5**
3.1 Failure time endpoint case . . . 6
3.1.1 Survival analyses . . . 6
3.1.2 Joint modeling of two failure time endpoints . . . 6
3.1.3 Assessment of Weibull and proportional hazard assumptions . . . 7
3.2 Marginal model with fixed effects (MFE) . . . 8
3.3 Bias in the estimation of measures of surrogacy . . . 8
3.4 Surrogate Threshold Effect . . . 9
3.5 Software . . . 10

4 Results **11**
4.1 Correlation between the two endpoints . . . 11
4.2 Hazard ratio for both endpoints per country . . . 11
4.3 Country specific Cox-proportional hazard model . . . 11
4.4 Joint distribution of two time endpoints . . . 13
4.5 Surrogate Threshold Effect . . . 14

5 Discussion and Conclusion **15**

6 Limitations and Recommendations **17**

Reference **19**

Appendix **23**

List of Figures

1 Distribution of patients per treatment per country 4
2 Kaplan-Meier curves for PFS (left) and OS (right) 4
3 Forest plots of hazard ratio on the PFS (left) and hazard ratio on OS (right).The vertical dashed line represents reference value one for hazard ratio. 11
4 Treatment effects on the true endpoint (OS) versus treatment effects on the surrogate endpoint (PFS) for all units of analysis. 12
5 Correlation between treatment effects on PFS and OS in historical trials. The sizes of the circles are proportional to the number of patients in each center or country. Predictions (short dashes) with 95%prediction limits leading to (long dashes) and (solid line) 14
6 Treatment effects on the true endpoint (overall survival) versus treatment effects on the surrogate endpoint (progression free survival) for all units of analysis from Clayton copula model 23
7 Treatment effects on the true endpoint (overall survival) versus treatment effects on the surrogate endpoint (progression free survival) for all units of analysis from Hougaard copula model 23
8 Treatment effects on the true endpoint (overall survival) versus treatment effects on the surrogate endpoint (progression free survival) for all units of analysis from Plackett copula model 24

List of Tables

1 Distribution of the patients and sites per country 3
2 Trial level surrogacy results 12
3 Results of the trial and individual-level surrogacy analysis 13

1 Introduction

Glioblastoma is the most frequent and most aggressive form of primary malignant brain tumor in adults. Median survival was generally less than one year from the time of diagnosis, and even in the most favorable situations, most patients died within two years. Standard therapy consisted of surgical resection to the extent that was safely feasible, followed by radiotherapy[38]. The addition of concomitant and adjuvant temozolomide (TMZ) to radiotherapy for newly diagnosed GBM resulted in clinically meaningful and statistically significant survival benefit with minimal additional toxicity[29].

Despite encouraging results from the recent ground breaking EORTC/NCIC trial, the median survival among patients with newly diagnosed GBM is 14.6 months with estimated 2-year survival of 27%. Stupp's trial changed the standard of care of these patients to include the combination of radiation therapy and TMZ followed by adjuvant TMZ.

With increasing research, more new molecular targeted agents are discovered and since there is no new standard care, it is imperative to develop more efficient clinical trials to evaluate these treatment strategies. Selecting one endpoint is a key factor to have a successful clinical trial. The investigator can make an efficient and accurate decision about the treatment efficacy only if the endpoint is timely and clinically meaningful. The most commonly used end points for phase II trials in patients with Glioblastoma multiform (GBM) is the proportion of patients who are alive and progression free at six months (six month progressive free-survival (PFS6) and the proportion of patients alive at 12-months (12-month overall survival (OS12)).

Overall survival (OS) is the primary endpoint of large adjuvant GBM phase III trials. OS12 can be accurately measured but may be confounded with subsequent therapies upon progression, where the converse is true for PFS6[25]. OS has the advantage of being the most objective measurement of efficacy. It is also easy to measure and the interpretation is straight forward. Though OS has the above mentioned advantages it also has the following limitations; OS requires a longer duration in this case 12 months from the time of last patient enrolled to the time of getting the results. Furthermore the effect of initial treatment on the OS is likely to be diluted by the administration of subsequent anticancer treatment after patients go off the study in question, typically due to disease progress, toxicity or completion of the trial treatment. When PFS6 is used as the primary endpoint, results can be obtained 6-months sooner than OS12 but it also has the following limitation; PFS6 is based on clinical or imaging criteria, both have an element of subjectivity and may be influenced by prior therapies[25,27]. Due to the limitations of OS listed above, PFS has been suggested in literature as a possible alternative primary endpoint for OS. It is therefore of interest to investigate whether PFS could replace OS as the primary endpoint in randomized GBM cancer trials.

A strong association was demonstrated between the endpoints PFS at 2, 4, and 6 and overall survival in the trial of 183 patients with newly diagnosed GBM enrolled in 3 phase II protocols at the university of California- san Francisco. This concluded that patients showed the signs of early progression were at significantly higher risk of earlier death and thus suggested the 6-month PFS may be appropriate primary end-point in the context of phase II up front GBM trials in the TMZ[27]. Another study which involved pooled data from 11 North Central Cancer Treatment Group trials for patients with newly diagnosed GBM(n=1348), they also found a strong association between progression free survival and overall survival. PFS6 seemed to be a reasonable end point for phase II trial in patients with recurrent GBM[30,31,32].

Potential surrogate endpoints may be laboratory variables (also called biomarkers) including biochemical markers, cellular markers cytokine markers, genetic makers, radiological measurements; physiological assessment or physical signs[13,30]. These results are likely to be expanded as the concept of surrogacy is becoming more widely adopted, even outside medical research. As the research on surrogate endpoints has evolved, potential surrogate have been extended from single measures to composite outcomes. This is particularly true in oncology clinical trials[32]. For example, disease free survival (DFS) and progression free survival (PFS) draw great attention in colorectal and breast cancers. These endpoints are a combination of time to death from any cause (OS) and time to disease recurrence and progression respectively.

In this project, two trials in newly diagnosed GBM were used where a standard treatment (surgery/biopsy followed by radiotherapy) is compared to a new treatment (surgery/biopsy followed by combined radiotherapy and chemotherapy). Two failure time endpoints used are PFS and OS. PFS was noted as time from randomization to disease progression or death from any cause while OS was time from randomization to death of any cause. Patients were censored at the date of last visit.

1.1 Objectives

The objective of this project is to evaluate whether PFS can be a surrogate for OS for patients with newly diagnosed GBM. The rest of this report is organized as follows. In chapters 2 we brief ly describe the data set used in this project. In chapter 3, we introduce the methods used to evaluate the surrogate endpoint. In chapter 4, we present the results, while chapter 5, we close with the discussion and make conclusion based on the results obtained.

2 Data Description

2.1 Trial data description

The data analyzed in this report is derived from two clinical trials conducted in sixteen countries namely; Australia, Austria, Belgium, France, Germany, Israel, Italy, Poland, Slovenia, Spain, Sweden, Switzerland, The Netherlands, United Kingdom, Canada and Hungary.
In the first trial (EORTC 26981-22981/NCCTG CE.3), 573 patients with newly diagnosed GBM according to local pathologists were enrolled in different hospitals in different countries. These patients were randomly assigned to radiotherapy alone (coded "0") or radiotherapy plus concomitant and adjuvant TMZ (coded "1"). The number of hospitals in a country ranged from 1 to 20 and the number of patients in each country ranged from 3 to 168. Table 1 shows the distribution of the patients and the hospitals per country. In the second clinical trial (EORTC 26882), there were 205 newly diagnosed GBM patients according to local pathologists who were randomized between radiotherapy alone (coded "0") and radiotherapy followed by adjuvant dibromodulcitol and BCNU chemotherapy (coded "1") randomized in different hospitals. There were only 7 countries involved with hospitals in each country ranging from 1 to 3 and patients in each country ranging from 5 to 51. The summary of the distribution of patients and hospitals per country is found in table 1.

Table 1: Distribution of the patients and sites per country

Trial	Country	Austr	Aust	BEL	F	GE	IS	PL	IT	SLV	SP	SD	SWI	NL	UK	CA	HU
	Hospitals	1	2	8	13	8	2	1	8	1	3	1	8	5	3	20	
EORT26981/22981	Patients	3	21	46	50	74	14	3	38	3	19	1	52	67	14	168	
	Hospitals			3	3	1							1	1	2		1
EORTC 26882	patients			33	31	5							25	17	51		43

In the first trial, proportion of censored survival time was 3.32% on PFS and 7.16% on OS. In the second trial, censoring proportion was 1.4% for PFS and 1.95% for OS.
In both trials, there was treatment effect on the PFS with P-value of <0.001 for the first trial and 0.0168 for the second trial using the log-rank test. On the OS, there was a treatment effect in first trial with p-value <0.001 while in the second trial, the treatment effect was non-significant with p-value of 0.0608 using the log-rank method.

2.2 Combined data description

Clinical data were obtained for 778 GBM patients after pooling the two datasets. The analysis was restricted to centers with at least three patients on each treatment arm. This constraint was adopted to ensure estimability of the joint copula models, as they required the estimation of six marginal parameters (two for the Weibull parameters for "overall survival", two for the Weibull parameters for "PFS", and two treatment effects) related to the marginal distribution

of surrogate and true endpoint for each trial i. In general, the minimum number of patients for estimability of the marginal parameters would require at least three patients per center, with at least one observed failure and at least one patient in each treatment group. As a result, data from 12 countries for a total of 768 patients were used in the analysis. 379 patients were assigned to radiotherapy alone and 389 patients were assigned to radiotherapy plus chemotherapy. Four countries were dropped from the analysis namely; Australia, Poland, Slovenia and Sweden since they had less than 6 patients or had no patients in some treatment arms.
Figure 1 below shows the distribution of patients per country per treatment. We can observe the patients were equally distributed to treatments and Canada had the highest number of patients overall.

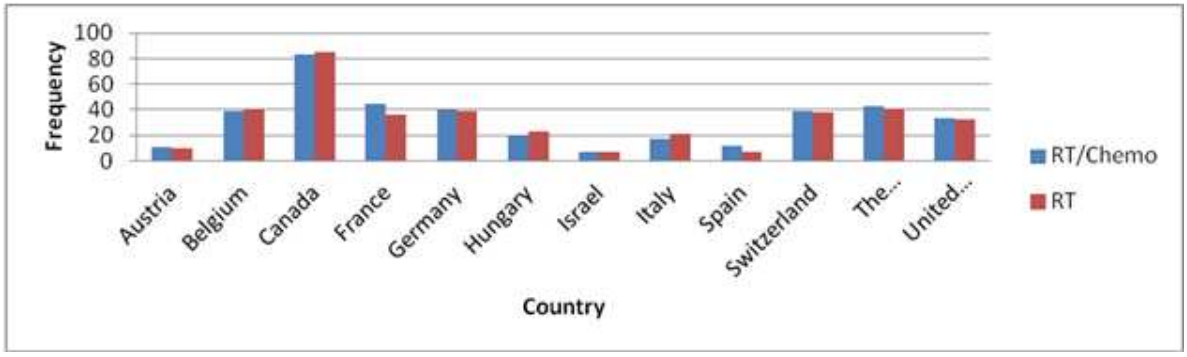

Figure 1: Distribution of patients per treatment per country

Prognostic factors like age ($0 =<= 50, 1 = 51 - 60, 2 => 60$), extent of surgery(0=complete resection,1=partial resection, and 2=biopsy) and use of steroids(0=No and 1=Yes) were collected. There were 135 patients with age 0, 413 patients with age 1 and 220 with age2.
746 PFS events (2.86% censored observations) and 723 OS events (5.86% censored observations) was observed. Figure 2 shows the Kaplan-Meier survival curves and we observed that overall survival has longer median survival time then progression free survival as expected. The treatment effect on both endpoints was significant with a P-value <0.0001.

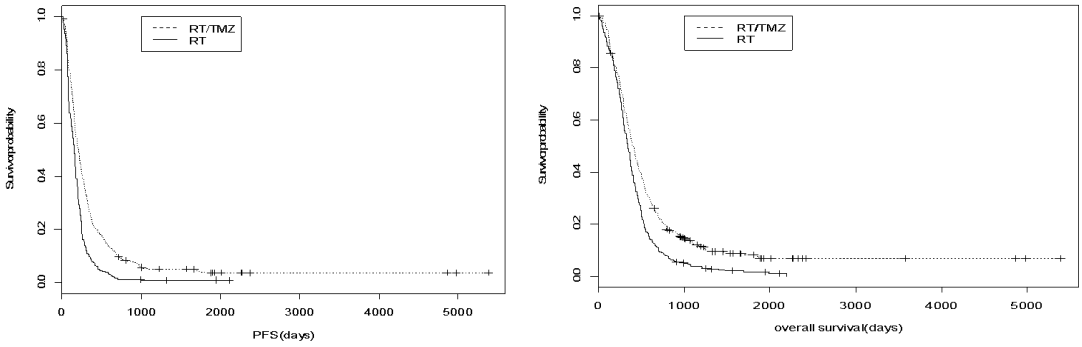

Figure 2: Kaplan-Meier curves for PFS (left) and OS (right)

3 Methodology

One of the most important factors influencing the duration and complexity of the process of developing new treatment is the choice of the endpoint which can be used to assess efficacy of the treatment. Two main criteria to select the endpoint are its sensitivity to detect treatment effects and its clinical relevance to goals of study[7,18].
Surrogate end points are very challenging to validate. The data should demonstrate that surrogate is prognostic for the true endpoint independently of treatment. The data should also show that treatment effect on the surrogate reliably predicts its effect on the true endpoint[12].
Prentice initially proposed that for a surrogacy relationship to be established the surrogate should predict the clinical end point, treatment should have a significant effect on both the candidate surrogate and the clinical end point. The treatment effect on the surrogate should capture the full effect of treatment on the clinical end point[12].

The surrogate can be validated using single trial method or meta-analytical validation methods. In the single trial setting, there are four operational criteria proposed by Prentice[7] and the associated estimate of surrogacy quantity (proportion of treatment effect explained by the surrogate (PTE))[34]. These methodologies and extensions were developed to reflect the spirit that a valid surrogate should fully capture the treatment effect on the clinical endpoint. However, there are many drawbacks of using single-trial data to validate or evaluate a putative surrogate endpoint. Prentice's criteria have been criticized as being too stringent to be applied in practice[7,28,34]. PTE methods rely on nontestable assumptions, require an extremely large sample size, and/or require an unrealistically highly significant treatment effect on the clinical points to obtain estimates with sufficient precision[7].

In order to generalize and make direct inference from the results we need to get data from different trials (multi-trial).This information collected from different trials will provide a better understanding of how well the treatment effect on the surrogate can predict the effect on the clinical endpoint[19].
To achieve validity of a surrogate endpoint, the candidate surrogate endpoint must be validated at both individual and trial level. Individual surrogacy (R^2_{ind}), measures the association between the potential surrogate endpoint and the clinical endpoint, adjusting for the effect of treatment across all the trials included[34]. This estimating involves joint modeling the surrogate and clinical endpoint. On the other hand, Trial surrogacy(R^2_{trial}), describes how well one can predict the treatment effect on the clinical endpoint in a future trial based on the observed association on the surrogate and clinical endpoint observed in previous trials[7,34]. Both R^2_{ind} and R^2_{trial} are coefficients of determination measures, and take values between zero and one. Values of and close to one indicate a stronger surrogacy than values close to zero. This clear and intuitive boundary provides a quantitative measure of surrogacy assessment[7,32,34].

In this chapter, in section 3.1 we discuss the meta-analysis based on failure-time endpoints where proportional hazard models were fitted using Cox models and copulas based on Weibull baseline hazard. Finally, in section 3.2 we discuss the method of surrogate threshold effect (STE)

3.1 Failure time endpoint case

3.1.1 Survival analyses

The analysis of the PFS and OS was conducted using intention-to-treat approach. The Kaplan-Meier method was used to estimate the distribution of both PFS and OS. Country was used as the unit of analysis. The treatment effect on both endpoints was estimated through a proportional hazard model with treatment as the only factor.

3.1.2 Joint modeling of two failure time endpoints

In this study both the surrogate and true endpoints are failure time variables, validation of surrogates in this setting is severely complicated by many factors, like the presence of censoring and competing risks, or absence of a unifying frame work such as the multivariate normal distribution[6].
To analyze this kind of data, Burzykowski proposed use of copula models [5,17,26,33]. They assumed that the joint survival function of (S_{ij}, T_{ij}) which is written as:

$$F(s,t) = P(S_{ij} \geq s, T_{ij} \geq t = C_\theta\{F_{S_{ij}}(s), F_{T_{ij}}(t)\}, s,t \geq 0 \tag{1}$$

where S_{ij} and T_{ij} are random variables denoting the surrogate and true endpoint respectively, for the jth subject in the ith center, while $F_{S_{ij}}$ and $F_{T_{ij}}$ denoted marginal survival functions and C_θ is a copula i.e. a bivariate distribution function on $[0,1]^2$ with uniform margins.
The attractive feature in model (1) is that the margins do not depend on the choice of the copula function. In principle, any copula function can be used in model (1). Here we considered three copula models namely Clayton copula, Hougaard copula and Plackett copula. We used proportional hazards model to model the treatment effect on the marginal distribution of S_{ij} and T_{ij} in (1) as proposed by Burzykowski

$$F_{S_{ij}}(s) = exp\Big\{ - \int_0^s \lambda_{S_i}(x) exp(\alpha_i Z_{ij}) dx \Big\} \tag{2}$$

$$F_{T_{ij}}(s) = exp\Big\{ - \int_0^t \lambda_{T_i}(x) exp(\beta_i Z_{ij}) dx \Big\} \tag{3}$$

where λ_{S_i} and λ_{T_i} are trial-specific marginal baseline hazard functions and α_i and β_i are trial-specific treatment Z effects on the endpoints in trial i[5].

The hazard function can either be specified parametrically or can be left unspecified as in the model proposed by Cox. In this study, copula models assuming a Weibull baseline hazard were fitted. The model fit has two stages: First stage, the parameters of marginal survival functions $F_{S_{ij}}$ and $F_{T_{ij}}$ estimated using proportional hazard model with a Weibull distribution for the marginal baseline hazards.
Second stage, Burzykowski proposed to use the model:

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \begin{pmatrix} a_i \\ b_i \end{pmatrix} \tag{4}$$

where the second term on the right hand side of (4) is assumed to follow a zero-mean normal distribution with dispersion matrix

$$D = \begin{pmatrix} d_{aa} & d_{ab} \\ d_{ab} & d_{bb} \end{pmatrix} \tag{5}$$

Using model (4) at the second stage of the two-stage approach, the quality of surrogate S at the trial level is assessed based on the coefficient of determination

$$R^2_{trial(r)} = \frac{d^2_{ab}}{d_{aa} d_{bb}} \tag{6}$$

Individual surrogacy was estimated by spearman's rho (ρs)[5]. Consider two random variables X and X_2 with marginal F_1 and F_2, respectively. Spearman's *rho* is defined by

$$\rho s = corr((F_1 X_1), (F_2 X_2))$$

and

$$corr(X_1, X_2) = \frac{cov(X_1, X_2)}{\sqrt{Var(X_1) Var(X_2)}}$$

The correlation fully describes the dependence structure.

3.1.3 Assessment of Weibull and proportional hazard assumptions

To assess the assumptions used in the copula of Weibull baseline distribution, the estimated log-cumulative hazard function was plotted against the logarithm of the survival time for different treatment groups. Parallel straight lines would indicate that both Weibull and proportional hazard assumption are suitable. If the lines are not particularly straight, the Weibull model may not be appropriate. However, if the lines can be taken as parallel, this would mean that the proportional hazard model is valid and Cox proportional hazard model can be applied.

3.2 Marginal model with fixed effects (MFE)

Cox proportional model was fitted separately to both endpoints in the first stage;

$$\lambda_{S_{ij}}(S_{ij}) = \lambda_{S_i} exp(\alpha_i Z_{ij}) \tag{7}$$

$$\lambda_{T_{ij}}(T_{ij}) = \lambda_{T_i} exp(\beta_i Z_{ij}) \tag{8}$$

where α_i and β_i are trial-specific treatment effects.
At the second stage the determination coefficients $R^2_{trial(r)}$ was computed from the regression of $\hat{\beta}_i$ on $\hat{\alpha}_i$.

3.3 Bias in the estimation of measures of surrogacy

Using model (2)-(3), the $R^2_{trial(r)}$ given by (6) was estimated by the square of the correlation coefficient between treatment effects α_i and β_i . As such, the square of the sample correlation coefficient is a biased estimator of coefficient of determination[7]. Several adjusted estimators have been proposed to reduce the bias.
Let us assume that the estimated treatment effect follow $\hat{\beta}_i$ on $\hat{\alpha}_i$ follow the model

$$\begin{pmatrix} \hat{\alpha}_i \\ \hat{\beta}_i \end{pmatrix} = \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} + \begin{pmatrix} \varepsilon_{a_i} \\ \varepsilon_{b_i} \end{pmatrix} \tag{9}$$

where the estimation errors ε_{a_i} and ε_{b_i} are normally distributed with mean zero and covariance matrix:

$$\Omega_i = \begin{pmatrix} \sigma_{aa,i} & \sigma_{ab,i} \\ \sigma_{ab,i} & \sigma_{bb,i} \end{pmatrix} \tag{10}$$

and $(\alpha_i, \beta_i)^T$ follow model (4) with matrix D given by (5). Consequently, $(\hat{\alpha}_i, \hat{\beta}_i)^T$ follows a normal distribution with mean $(\alpha, \beta)^T$ and matrix $D + \Omega_i$.

In this study, we considered the method developed by Fuller, for measurement errors models with error in the equation and unequal error variances. Assuming the true unobserved trial-specific treatment effects follow the simple linear regression model

$$\beta_i = \gamma_0 + \gamma_1 \alpha_i + \varepsilon_i, \tag{11}$$

where γ_0 and γ_1 are constant coefficients and ε_i is a random variable with mean 0 and variance σ. Let the estimates $\hat{\beta}_i$ and $\hat{\alpha}_i$ follow model (11). We will also assume that ε_i is independent of $(\varepsilon_{a_i}, \varepsilon_{b_i})$.
Then we can write trial surrogacy as

$$R^2_{trial(r))} = \frac{\gamma_1^2 d_{aa}}{\gamma_1^2 d_{aa} + \sigma} \tag{12}$$

The estimator of $R^2_{trial(r))}$ based on formula(12) is easier to compute than the one based on the maximum likelihood estimate of the matrix D, say, obtained by fitting the model defined by (9)-(10) and (4)-(5) to the estimated pairs $(\hat{\beta}_i, \hat{\alpha}_i)$[7].

3.4 Surrogate Threshold Effect

The validity of a surrogate is quantified by the coefficient of determination $R^2_{trial(r))}$ obtained from (6) - (12), which allows for prediction of the treatment effect on the true endpoint given the observed treatment effect on the surrogate. One problem related to the use of $R^2_{trial(r))}$ is the difficulty in interpreting its value. To address this problem, Burzykowski proposed new concept called surrogate threshold effect (STE), defined as the minimum treatment effect on the surrogate required to predict a non-zero treatment effect on the true endpoint in future trial[7]. The magnitudes of the STE reflect the minimum width of the prediction limits for the treatment effect on OS in a new trial, obtained from the effect on PFS. The smaller the STE, the narrower the prediction limits, and the more useful the surrogate[7].
One of its interesting features, apart from providing information relevant to the practical use of a surrogate endpoint, is its natural interpretation from a clinical point of view.
There are two versions of STE. The first is denoted as $STE_{\infty,\infty}$ and is obtained using the following predictive variance:

$$\sigma = d_{bb}(1 - R^2_{trial(r)}) \tag{13}$$

$$STE_{\infty,\infty} = \alpha - \frac{d_{aa}}{d_{ab}}\Big\{\beta + Z_{1-\frac{\gamma}{2}}\sqrt{d_{bb}(1 - R^2_{trial(r)})}\Big\} \tag{14}$$

where σ is the residual variance, α and β are treatment effect on the surrogate (PFS) and true endpoint (OS) respectively. $Z_{1-\frac{\gamma}{2}}$ is the $\mathbf{1} - \frac{\gamma}{2}$ quantile of standard normal distribution. The infinity signs used in the notation for $STE_{\infty,\infty}$ indicates that (14) assumes the knowledge of parameter vector $\upsilon \equiv (\beta, \alpha, d_{aa}, d_{ab}, d_{bb})$. This is achievable only with an infinite number of infinite-sample-size trials in the meta-analysis data and infinite sample size of the new trial. The second is denoted by $STE_{N,\infty}$ with N indicating the need for estimation of υ[7].

Two-stage modeling strategy was used, where in the first stage a joint model of S and T is fitted to meta-analytic data. This model (7)-(8) provided estimates $\hat{\beta}_i$ and $\hat{\alpha}_i$ of the trial-specific treatment effects $\beta_i \equiv \beta + b_i$ and $\alpha_i \equiv \alpha + a_i$. At stage two, a model was fitted to $\hat{\beta}_i$ and $\hat{\alpha}_i$, allowing the estimation of parameter vector υ[7].
Now, the estimate of υ and its variance-covariance matrix could be used to compute the prediction variance. We can assume that the trial-specific treatment effect $\hat{\beta}_i$ and $\hat{\alpha}_i$ follow the simple linear regression model (11)
Model (11) was fitted to estimated treatment effect with and without the adjustment for the estimation error in the treatment effect. For the adjusted model, the prediction limits defining $STE_{\infty,\infty}$ and $STE_{N,\infty}$ were computed using

$$\tilde{\sigma} = \frac{1}{N-2}\sum_{i=1}^{N}\Big(\hat{\beta}_i - \hat{\gamma}_0 - \hat{\gamma}_1\hat{\alpha}_i\Big) - \sum_{i=1}^{N}\Big(\hat{\sigma}_{bb,i} - 2\tilde{\gamma}_1\hat{\sigma}_{ab,i} - 2\tilde{\gamma}_1^2\hat{\sigma}_{aa,i}\Big), \tag{15}$$

$$var(\hat{\beta}) = \sigma + var(\tilde{\gamma}_0) + 2\alpha cov(\tilde{\gamma}_0, \tilde{\gamma}_1) + (\alpha^2 + d_{aa})var(\tilde{\gamma}_1), \tag{16}$$

Respectively, the variance-covariance matrix of the estimators $\tilde{\gamma}_0$ and $\tilde{\gamma}_1$ can be obtained using a robust estimator. The prediction limits for the unadjusted model were computed using the formulas for a simple linear model. $STE_{\infty,\infty}$ limits were obtained using the mean residual sum of squares for the regression of $\hat{\beta}_i$ and $\hat{\alpha}_i$ as the estimate of prediction variance σ:

$$\hat{\sigma} = \frac{1}{N-2} \sum_{i=1}^{N} \left(\hat{\beta}_i - \hat{\gamma}_0 - \hat{\gamma}_1 \hat{\alpha}_i \right)^2 \tag{17}$$

Whereas for $STE_{\infty,\infty}$ the prediction variance was estimated by

$$var(\hat{\beta}) = var(\hat{\gamma}_0) + var(\hat{\gamma}_1) + cov(\hat{\gamma}_0, \hat{\gamma}_1) + \hat{\sigma} \tag{18}$$

Where $var(\hat{\gamma}_0)$ and $var(\hat{\gamma}_1)$ are variance of the intercept and slope, respectively while $cov(\hat{\gamma}_0, \hat{\gamma}_1)$ is the covariance got from fitting model (11) and $\hat{\sigma}$ is the mean residual sum of squares.

3.5 Software

In this report SAS MACRO using IML programming language, developed by Burzykowski were used to estimate the copula. The other analyses were conducted in SAS version 9.3 and R 2.13.1.

The level of significance used in this report is 5%.

4 Results

4.1 Correlation between the two endpoints

The number of events observed for PFS (746) and for OS (723) was similar. PFS and OS were highly correlated with a rank correlation of 0.845.

4.2 Hazard ratio for both endpoints per country

Figure 3 presents the forest plot for hazard ratio obtained on PFS and OS. For the PFS we observed that most confidence intervals did not contain one meaning there was a treatment effect except for Israel, Italy Austria Switzerland and Hungary. On the other hand, more intervals on OS contained one suggesting no significant treatment effect. Italy showed no treatment effect on progression free survival but we observed a treatment effect on overall survival. Belgium, United Kingdom and France showed treatment effect on PFS but no treatment effect on the OS.

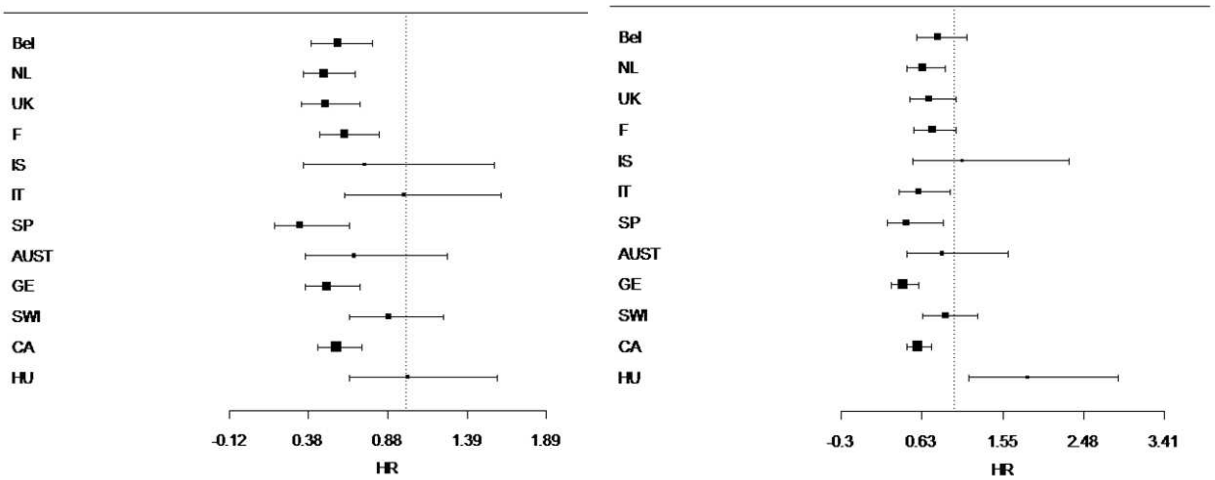

Figure 3: Forest plots of hazard ratio on the PFS (left) and hazard ratio on OS (right).The vertical dashed line represents reference value one for hazard ratio.

4.3 Country specific Cox-proportional hazard model

As mentioned in the chapter 3, a Cox's regression model was fitted to both endpoints and treatment effect was estimated per country. The log hazard ratio on the OS was regressed on the log hazard estimates from PFS. Table 2, shows the results obtained from the linear regression models both weighted and unweighted regression. The coefficient of determination for the unweighted regression was 0.453 while the value for the weighted regression was 0.429. The coefficients of determination were more or less the same for both the weighted and unweighted regression implying that weighting had a very small impact on the regression.

To investigate whether taking into account important prognostic factors like age and extent of surgery would change results of the analysis; the linear regression was refitted with age and extent of surgery as covariates in marginal models (7)-(8). The estimates of R-squared as presented in Table 2 are slightly increased, as compared estimates from models with only treatment as the only covariate.

Table 2: Trial level surrogacy results

Regression model	**Unweighted**	**Weighted**
R-squared without covariates	0.453(0.036, 0.86)	0.429(0.062,0.852)
R-squared with covariates	0.4917(0.088, 0.89)	0.5092(0.1128, 0.9055)

Figure 4 shows a plot of treatment effects on true endpoint (logarithm of HR OS) by treatment effects on the surrogate (logarithm of HR PFS), corresponding to regression model for both weighted and unweighted. The effects are weakly correlated. The results in table 3 confirm these observations.

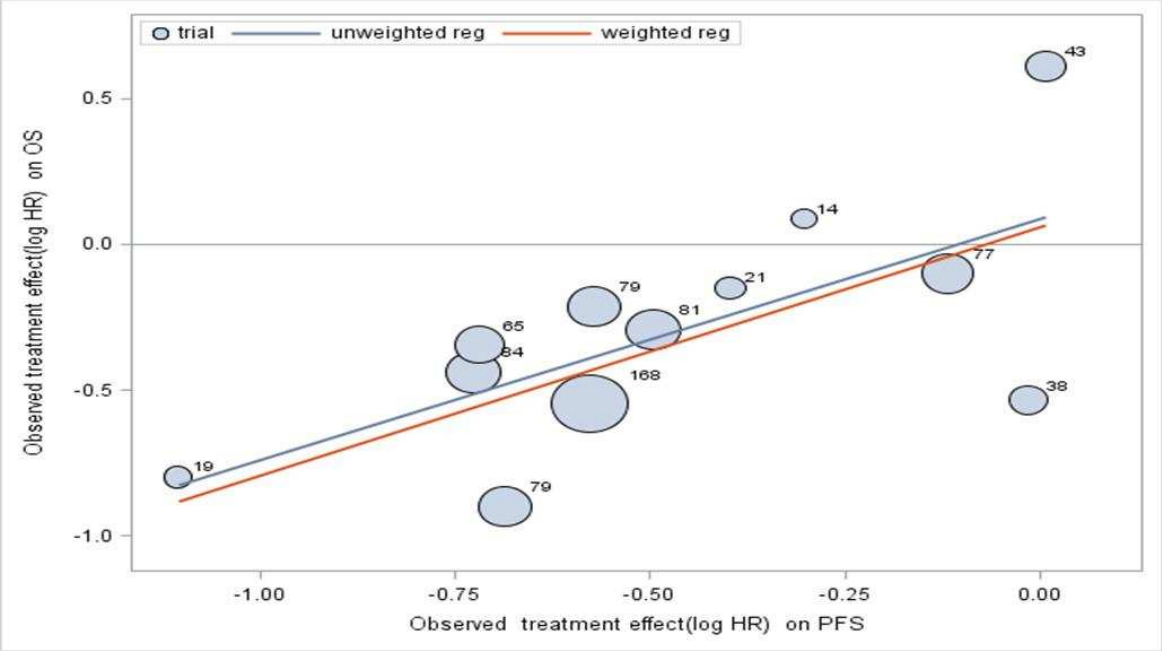

Figure 4: Treatment effects on the true endpoint (OS) versus treatment effects on the surrogate endpoint (PFS) for all units of analysis.

4.4 Joint distribution of two time endpoints

Table 3 presents results from the copula models. For all the models, two values of $R^2_{trial(r)}$ were presented unadjusted and adjusted. The unadjusted trial-level surrogacy was obtained by computing the correlation coefficient of pairs $(\hat{\alpha}_i, \hat{\beta}_i)$ without adjusting for measurement error in $\hat{\alpha}_i$ and $\hat{\beta}_i$). The unadjusted estimate suggests values of $R^2_{trial(r)}$ range from 0.481 to 0.775. For the estimates of $R^2_{trial(r)}$ adjusted for measurement errors ranges from 0.896 to 0.996. The Spearman's *rho* estimates range from 0.64 to 0.82. The highest likelihood was obtained for Plackett copula and lowest was for Clayton copula. This suggests that Plackett copula was the best of the three models for this data.

We will use the results obtained in table 3, for Plackett copula. The value for unadjusted $R^2_{trial(r)}$ was 0.481 (95% confidence interval [0.0746,0.888]). Adjusted $R^2_{trial(r)}$ was 0.896 (95% confidence interval [0.559,1.233]).

Table 3: Results of the trial and individual-level surrogacy analysis

Model	Likelihood	Individual	Trial(unadjusted)	Trial(adjusted)
Clayton	6.4	0.748(0.745,0.75)	0.775(0.552,0.99)	-
Hougaad	19.3	0.821(0.82,0.81)	0.734(0.476,0.991)	0.996(0.735,1.257)
Plackett	25.6	0.642(0.635,0.64)	0.481(0.0746,0.888)	0.896(0.559,1.233)

The results from unadjusted $R^2_{trial(r)}$ from Plackett copula were close those obtained from a two stage method. Although the confidence intervals for adjusted $R^2_{trial(r)}$ are wide, we can say that adjusted $R^2_{trial(r)}$ is at least 0.559 for the Plackett copula. The figure (6,7 and 8) in the appendix show plots of treatment effects on the true endpoint(logarithm of HR OS) by the treatment effects on the surrogate endpoint(logarithm of HR PFS), corresponding to the three copula models considered in the analysis. The size of the circles is proportional to the number of patients in each center or country. Figures 6 and 7 in the appendix show a moderate correlation between the treatment effects of PFS and OS. Figure 8 in the appendix show a weak correlation between the treatments of PFS and OS. An indication that the effects were weakly correlated. The results shown in the Table 3 confirm this.

4.5 Surrogate Threshold Effect

Using results obtained from table 2, for country specific Cox proportional hazard. Mean treatment effect on true endpoint and surrogate endpoint were equal to -0.52 and -0.39 respectively. In the analysis unadjusted for measurement errors, the estimate for $R^2_{trial(r)}$ was 0.453 (0.036, 0.869). Using simple regression model (11), γ_0 and γ_1 with corresponding standard errors were estimated by $\hat{\gamma}_0$ equal to 0.089 (s.e. 0.163.) and $\hat{\gamma}_1$ equal to 0.8269(s.e. 0.287) respectively. These estimates led to $STE_{\infty,\infty} = -0.789$ and $STE_{N,\infty} = -1.725$. Note that $STE_{N,\infty}$ and $STE_{\infty,\infty}$ were computed using the upper limits. The negative values of treatment effect, pointing to a reduction of the risk of failure, this was considered beneficial. Figure 5, shows the correlation between treatment effects on PFS and OS .The $STE_{N,\infty}$ was (on log-hazard scale) -1.725, which corresponds to PFS hazard ratio of 0.178. Thus, in order to predict a non-zero treatment effect on OS in a future trial, a hazard ratio of at most 0.178 or less needs to be observed on PFS. The value of STE was not close to the treatment effects on the surrogate endpoint observed in the meta-analysis. This suggests that PFS may not be considered as a surrogate for OS.

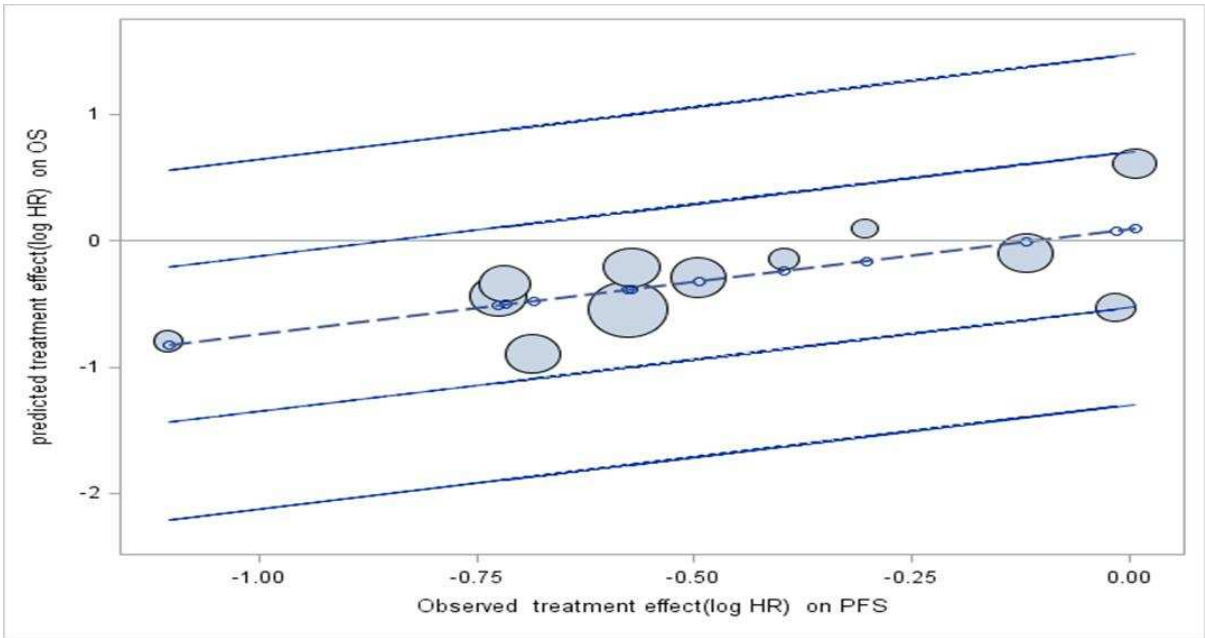

Figure 5: Correlation between treatment effects on PFS and OS in historical trials. The sizes of the circles are proportional to the number of patients in each center or country. Predictions (short dashes) with 95%prediction limits leading to (long dashes) and (solid line)

5 Discussion and Conclusion

The main objective of this project was evaluation of PFS as a surrogate for OS of patients newly diagnosed with GBM. Data from two clinical trials was used in this project. Sixteen countries participated in the trials. The analysis was restricted to centers with at least 3 patients on each treatment arm. This constraint was adopted to ensure estimability of the joint copula models, as they required the estimation of six marginal parameters (two for the Weibull parameters for "overall survival", two for the Weibull parameters for "PFS", and two treatment effects) related to the marginal distribution of surrogate and true endpoint for each trial i.

To get insight of the dataset, exploratory data analysis was done. Kaplan Meier survival curves showed longer median survival time in the OS as compared to PFS. These results were expected since PFS is an intermediate end point while overall survival is a late end point with respect to time of occurrence. Treatment effect was observed on both endpoints with a p-value <0.001.A high correlation was observed between PFS and OS (0.845). This observation alone does not make PFS a good surrogate for OS. Although there is no consensus regarding the theoretical conditions required for a surrogate endpoint to be valid, recent works suggests that surrogacy can be assessed through correlation between the end points and the treatment effects on these endpoints in a series of trials[8,9].

Cox proportional hazard model and copula models based on Weibull baseline distribution were used to obtain the measure of individual and trial surrogacy since both endpoints were failure-time. In the two-stage, Cox proportional hazard model was fitted in the first stage and simple regression in the second stage. The estimate for unadjusted $R^2_{trial(r)}$ (0.45) shows insufficient evidence to consider PFS a valid surrogate for OS in GBM. The adjusted $R^2_{trial(r)}$ obtained was outside the boundary of [0, 1]. In such a case this estimate is taken as non-defined. Such problem can be expected since none of the parameters in equation (12) is guaranteed to be positive[17]. The problem may be due to large magnitude of the measurement error present in the estimates of treatment effects. It was also interesting to check if accounting for important prognostic factors like age and extent of surgery would affect the results .The results show a slight increase in the estimate of R^2 but the result are not any different from the ones without covariates.

Plackett copula model was chosen as the best fit for the data since it had the highest likelihood. The Plackett copula results in table 3 also shows insufficient evidence to consider PFS a valid surrogate for OS. The adjusted $R^2_{trial(r)}$ from Plackett copula model suggested that adjusted $R^2_{trial(r)}$ was at least 0.559. This result still did not give sufficient evidence to consider PFS a valid surrogate for overall survival.

The results obtained from univariate Cox proportional model and Plackett copula were close. These results also reflect what was observed in exploratory analysis (Fig 3). Some centers showed treatment effect on the PFS while no treatment effect on the OS and the reverse was true. This suggests that PFS can not quite predicting treatment effect on the OS. The results from the copula models should be interpreted with caution since the assumption of Weibull baseline distribution was not suitable for this dataset as observed in the exploratory analysis.

In this report, we extended the validation methodology to investigate the predictive value of PFS as surrogate endpoint for OS. The surrogate threshold effect (STE) was used to validate surrogacy. $STE_{N,\infty}$ and $STE_{\infty,\infty}$ can be used to address the concern about the usefulness of meta-analytic approach to the validation of surrogate endpoints expressed by Gail[18]. The estimated value for surrogate threshold effect ($STE_{N,\infty}$) was -1.725 corresponding to hazard ratio of 0.178. This shows that if treatment ascertains a hazard ratio of 0.178 or less, it would be expected to achieve a non-zero treatment effect on OS. This suggests need of a large treatment effect to be observed on PFS. This high treatment effect has not been achieved by any clinical trial yet. From the meta-analysis, the treatment effect observed on the PFS in this project was -0.52 corresponding to hazard ratio of 0.6. This hazard ratio of 0.6 observed on PFS is higher than 0.178 of STE. This clearly suggests that PFS would not be a good surrogate for OS. This leads to the conclusion that PFS is not a valid surrogate for OS for GBM.

In conclusion, the results suggest that PFS is not a valid surrogate for OS neither at trial-level nor individual-level. It should probably not be used as a surrogate for OS in GBM for treatments of the type used in the trials analyzed. Furthermore, the surrogate threshold effect was too large compared to the treatment effect observed on PSF using meta-analysis. PFS does not convey clinically useful surrogacy properties for OS.

6 Limitations and Recommendations

This project was limited to only two trials which involved a small number of countries with limited sample size. Results for surrogate threshold effect were obtained without accounting for measurement errors. The results may be interpreted with caution. I would recommend further analysis could be repeated and should incorporate other large GBM phase III data like AVAGLIO, EORTC CENTRIC+CORE, and RTOG 0525 trials. This would enable the estimation errors to be accounted for in the construction of the prediction limits for the OS hazard ratio.

.

Reference

1 Bleehen, N.M., Newlands, E.S., Lee, S.M., Thatcher, N., Selby, P., Calvert, A.H., Rustin, G.J.S., Brampton, M. and Stevens, M.F.G. (1995) Cancer research campaign phase II trial of temozolomide in metastatic melanoma. *Journal of clinical oncology* **13**: 910-913

2 Bower, M., Newlands, E.S., Bleehen, N.M., Brada, M., Begent, R.J.H., Calvert, H., Colquhoun, I., Lewis, P. and Brampton, M.H. (1997) Multicentre CRC phase II trial of temozolomide in recurrent or progressive high-grade Glioblastoma. *Cancer chemotherapy and pharmacology* **40**: 484-488

3 Brock, C.S., Matthews, J.C., Brown, G., Newlands, E.S., and Price, P. (1997) In vivo demonstration of 11C-temozolomide uptake by human recurrent high grade astrocytomas. *British Journal of Cancer* **75**: 1241

4 Buckner, J.C. (2003) Factors influencing survival in high-grade Glioblastoma. *Seminar in Oncology* 30: Suppl **19**: 10-4.

5 Burzykowski, T., Molenberghs, G., Buyse, M., Renard, D., and Geys, H. (2001) Validation of surrogate endpoints in multiple randomized clinical trials with failure-time endpoints. *Applied Statistics*, **50**: 405-422.

6 Burzykowski, T. and Jose, C.A. (2005) Validation in the Case of Two Failure-time Endpoints. *Statistics for Biology and Health* **11**: 163-194

7 Burzykowski, T., Molenberghs, G., and Buyse, M. (2005) *The Evaluation of Surrogate Endpoints*. Statistics for Biology and Health: Springer

8 Burzykowski, T., and Buyse, M. (2006) Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. *Pharmaceutical Statistics* **5(3)**:173-86.

9 Buyse, M., Molenberghs, G. (1998) Criteria for the validation of surrogate endpoints in randomized experiments. *Biometrics* **54**: 1515-1527

10 Buyse, M., Molenberghs, G., Burzykowski, T., Renard, D., Geys, H. (2000) The validation of surrogate endpoints in meta-analyses of randomized experiments. *Biostatistics* **1**: 49-67

11 Buyse, M., Burzykowski, T., Carroll, K., Michiels, S., Daniel, J., Miller, L., and Piedbois, P. (2007) Progression-Free Survival is a Surrogate for Survival in Advanced Colorectal Cancer. *Journal of Clinical Oncology* **20**: 33

12 Buyse, M., Sargent , D.J., Grothey, A., Matheson , A., and Gramont , A. (2010) Biomarkers and surrogate end points the challenge of statistical validation; *Nature Reviews Clinical Oncology* **7**309-317

13 Chang, C.H., Horton, J., and Schoenfeld, D. (1983) Comparison of postoperative radiotherapy and combined postoperative radiotherapy and chemotherapy in the multidisciplinary management of malignant Gliomas: A joint Radiation Therapy Oncology Group and Eastern Cooperative Oncology Group study. *Cancer*; **52**: 997-1007.

14 Cox, D.R. (1972a) Regression models and life-tables. *Journal of the Royal Statistical Society* Series B:**34**, 187-202

15 Curran, W.J.Jr., Scott, C.B., Horton, J., Nelson, J.S., Weinstein, A.S., Fischbach, A.J., Chang, C.H., Rotman, M., Asbell, S.O., and Krisch, R.E. (1993) Recursive partitioning analysis of prognostic factors in three Radiation Therapy Oncology Group malignant glioma trials. *Journal of the National cancer institute* **85**: 704-10. 3.

16 De Gruttola, V.G., Clax, P., and DeMets, D.L. (2001) Considerations in the evaluation of surrogate endpoints in clinical trials: summary of National Institute of Health Workshop. *Control Clinical Trials* **22**: 485-502

17 Fleming, T.R. and Demets, D.L. (1994) Surrogate endpoint in clinical trials: are we being misled? *Annals of international Medicine* **125**: 605-613

18 Fleming, T.R. (1996) Surrogate endpoint in clinical trials, *Drug Information Journal* **30**, 545-551

19 Fuller, W.A. (1987) *Measurement Error Models.* New York: John Wiley & Sons

20 Gail, M.H., Pfeiffer, R., Van Houwelingen, H.C., and Carroll, R.J. (2000) On meta-analytic assessment of Surrogate outcomes. *Biostatistics*, **1**: 231-246.

21 Genest, C. and McKay, J. (1986) The joy of copulas: bivariate distributions with uniform marginal. *American Statistician* **40**: 280-283.

22 Green, S.B., Byar, D. P., Walker, M.D., Pistenmaa, D.A., Alexander, E.Jr., Batzdorf, U., Brooks, W,H,, Hunt, W.E., Mealey, J.Jr., Odom. G.L., Paoletti, P., Ransohoff, J2nd., Robertson, J.T., Selker, R.G., Shapiro, W.R., Smith, K.R. Jr., Wilson, C.B., Strike,T.A.(1983). Comparisons of carmustine, procarbazine, and high-dose methylprednisolone as additions to surgery and radiotherapy for the treatment of malignant glioma. *Cancer treatment reports* **67**: 121-32.

23 Hughes, M.D. (2002) Evaluating surrogate endpoints. *Control Clinical Trials* **23**: 703-707

24 Johnston, K. (1999) what are surrogate outcome measures and why do they fail in clinical research? *Neuroepidemiology* **18**: 167-173

25 Karla, V.B., Jan, C.B., Paul, D.B., Caterina, G., Patrick, J. F., Betsy, R.L, and Kurt, A.J. (2007) The relationship between six month progression- free survival and 12-months overall survival endpoint for phase II trials in patients with Glioblastoma multiforme. *Neuro-Oncology* **9**: 29-38

26 Lin, D.Y., Fleming, T.R., De Gruttola, V. (1997) Estimating the proportion of treatment effect explained by a surrogate marker. *Statistics in Medicine* **16**: 1515-1527

27 Mei-Yin, C., Polley, K.R., Nicholas, B., Jennifer, L.C, Michael, P. (2009) six-month progression- free survival as an alternative primary efficacy endpoint at overall survival in newly diagnosed Glioblastoma patients receiving Temozolomide. *Neuro-Oncology* **12(3)**: 274-282, 2010.

28 Molenberghs, G., Buyse, M., Geys, H. (2002) Statistical challenges in the evaluation of surrogate endpoints in randomized trials. *Control Clinical Trials* **23**: 607-625

29 Nelsen, R.G. (1999) An introduction to copulas. Lecture Notes in Statistics, 139. New York: Springer-Verlag.

30 Newlands, E.S., Stevens, M.F.G., Wedge, S.R., Wheelhouse, R.T., and Brock, C. (1997) Temozolomide: a review of its discovery, chemical properties, pre-clinical development and clinical trials. *Cancer treatment report* **23**: 35-61.

31 Patel, M., McCully, C., Godwin, K. and Balis, F. (1995) Plasma and cerebrospinal fluid pharmacokinetics of temozolomide. *The American Society of Clinical Oncology* **14**: 461 [Abstract 1485]

32 Roger Stupp, M.D., Warren P. Mason, MD., Martin, J. van den Bent, M.D., Michael Weller, Barbara Fisher, M.D., Martin J.B. Taphoorn, M.D., Karl, B., M.D., Alba A. Brandes, M.D., Christine Marosi, M.D., Ulrich Bogdahn, M.D., JÃ$\frac{1}{4}$rgen Curschmann, M.D., Robert, C. Janzer, M.D., Samuel K. Ludwin, M.D.,Thierry, G., Anouk, A., Denis, L., Gregory, C., Elizabeth, E., RenÃ©, O. (2005) Radiotherapy plus Concomitant and Adjuvant Temozolomide for Newly Diagnosed Glioblastoma. *The new England journal of Medicine* **352**: 10

33 Shih, J.H. and Louis, T.A. (1995a) Inferences on association parameter in copula models for bivariate survival data. *Biometrics* **51**: 1384-1399.

34 Shi, Q. and Sargent, D.J. (2009) Meta-analysis for the evaluation of surrogate endpoints in cancer clinical trials. *The Japan Society of Clinical Oncology* **14**: 102-111.

35 Stewart, L. A. (2002) Chemotherapy in adult highgrade glioma: a systematic review and Meta analysis of individual patient data from 12 randomised trials. *Lancet* **359**: 1011-8.

36 Tibaldi, F.S., Cortias , J., Molenberghs, G., Renard, D., Burzykowski, T., Buyse, M., Parmar, M., Stijnen, T., and Wolfnger, R. (2003) Simplified hierarchical linear models for the evaluation of surrogate endpoints. *Journal of Statistical Computing and Simulation* **73**: 643-658.

37 Van Houwelingen, J.C., Arends, L.A., and Stijnen, T. (2002) Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in Medicine* **21**: 589-624.

38 Walker, M.D., Alexander, E., and Hunt, W.E. (1978) Evaluation of BCNU and/or radiotherapy in the treatment of anaplastic gliomas: a cooperative clinical trial. *Journal of neurosurgery* **49**: 333-43

39 Walker, M.D, Green, S.B., Byar, D.P., Batzdorf, U., Brooks, W.H., Hunt, W.E., MacCarty, C.S., Mahaley, M.S.Jr., Mealey, J.Jr., Owens, G., Ransohoff, J. 2nd.,Robertson, J,T,, Shapiro, W,R,, Smith, K,R, Jr,, Wilson, C.B., Strike, T,A. (1980) Randomized comparisons of radiotherapy and nitrosoureas for the treatment of malignant glioma after surgery. *The new England journal of Medicine* **303**: 1323-9.

40 Yung, W.K., Albright, R.E., Olson, J., (2000) A phase II study of temozolomide vs. procarbazine in patients with Glioblastoma multiforme at first relapse. *British Journal of Cancer* **83**: 588-93.

Appendix

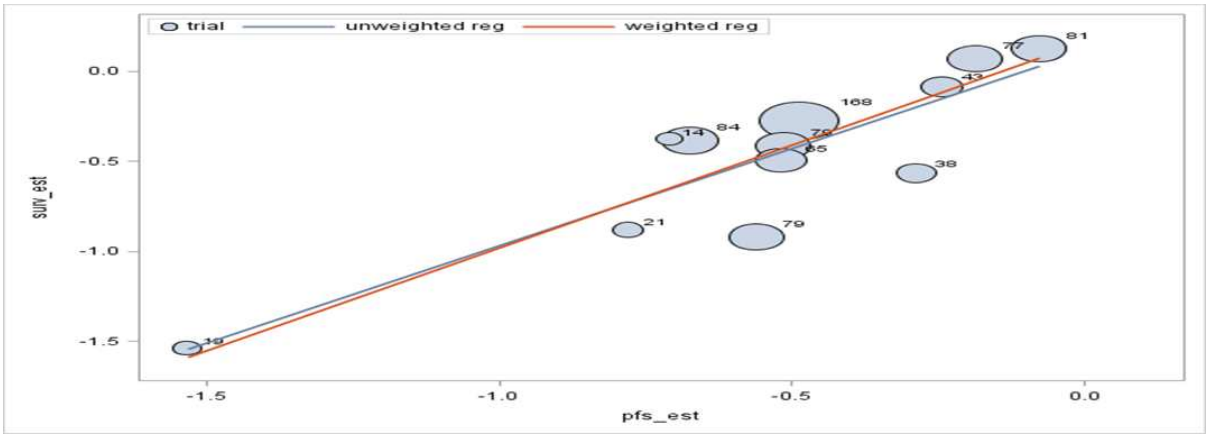

Figure 6: Treatment effects on the true endpoint (overall survival) versus treatment effects on the surrogate endpoint (progression free survival) for all units of analysis from Clayton copula model

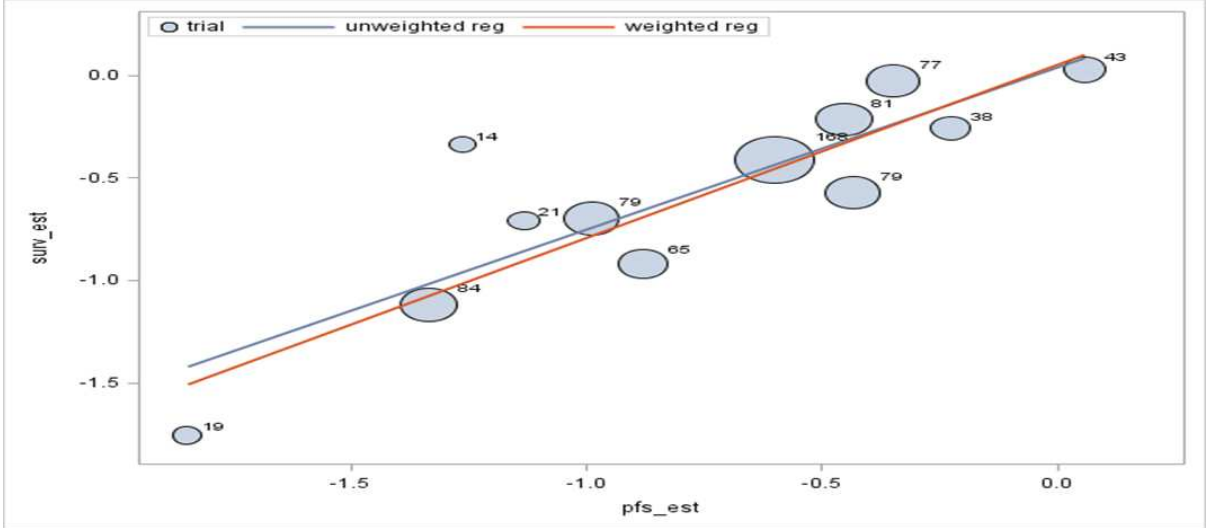

Figure 7: Treatment effects on the true endpoint (overall survival) versus treatment effects on the surrogate endpoint (progression free survival) for all units of analysis from Hougaard copula model

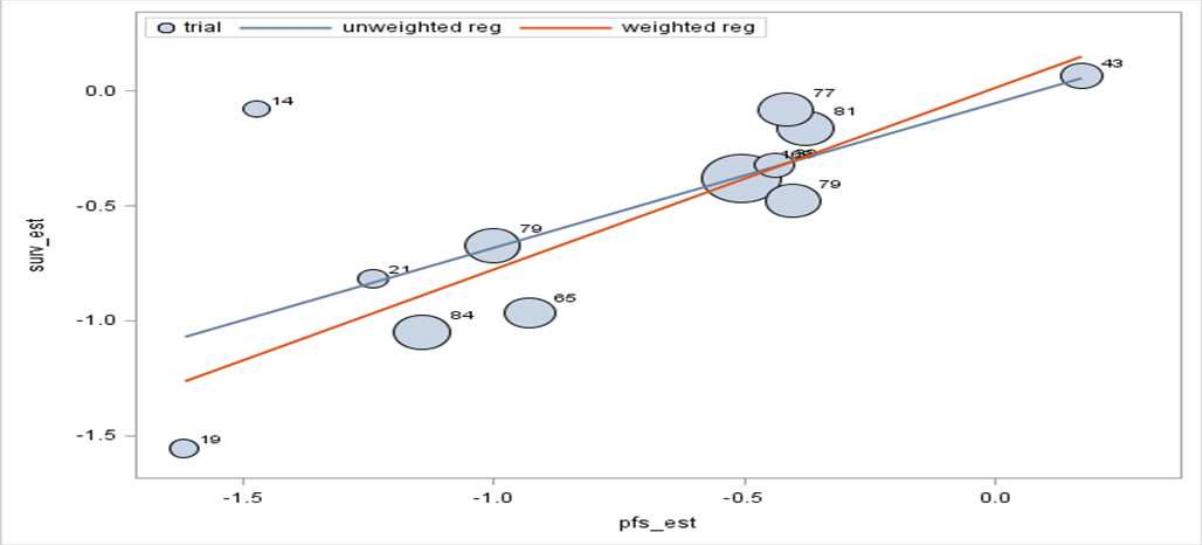

Figure 8: Treatment effects on the true endpoint (overall survival) versus treatment effects on the surrogate endpoint (progression free survival) for all units of analysis from Plackett copula model

Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
Validation of PFS as surrogate for overall survival for GBM

Richting: **Master of Statistics-Biostatistics**
Jaar: **2013**

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Voor akkoord,

Nangosyah, Julius

Datum: **11/09/2013**